

The Geography of Linguistic Diversity and the Provision of Public Goods

ONLINE APPENDIX

Klaus Desmet
SMU, NBER and CEPR

Joseph Gomes
UNAV

Ignacio Ortuno
UC3M

October 2018

Table 1: Global ELF and Local-Global Ethnolinguistic Complementarity – No GDP

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Child Survival	Measles Immunization	Hospital Beds	Literacy Rate	Schooling	Improved Sanitation	Road Density
Global ELF	-7.089*** (1.362)	-27.226*** (4.590)	-1.957*** (0.651)	-39.867*** (5.737)	-0.446*** (0.158)	-43.155*** (8.134)	-2.154 (1.976)
Local-Global Compl.	20.512*** (4.996)	73.608*** (15.949)	8.867*** (2.794)	131.210*** (19.538)	1.887*** (0.489)	136.707*** (29.971)	13.596 (11.577)
Absolute Latitude	0.068*** (0.021)	0.237*** (0.082)	0.114*** (0.017)	0.323** (0.124)	0.013*** (0.003)	0.462** (0.182)	0.250*** (0.068)
Latin America & Carib.	1.037 (0.669)	4.941* (2.741)	0.537 (0.466)	7.219* (3.788)	0.250** (0.100)	0.230 (5.925)	3.089* (1.777)
Sub-Saharan Africa	-7.264*** (0.878)	-10.723*** (2.839)	0.054 (0.463)	-13.451*** (3.982)	-0.241** (0.115)	-34.361*** (6.111)	1.406 (1.952)
East and S.E. Asia	0.445 (0.980)	2.410 (3.680)	1.444 (1.024)	12.101** (5.338)	0.062 (0.119)	-6.944 (8.415)	-3.276 (2.331)
French Legal Origin	0.224 (0.789)	-0.109 (2.918)	-0.431 (1.512)	-4.401 (6.561)	0.022 (0.099)	0.735 (5.341)	-10.799* (6.307)
German Legal Origin	0.977 (0.690)	1.228 (2.432)	1.325 (1.604)	-0.171 (6.800)	0.200** (0.090)	2.512 (5.606)	-8.374 (6.404)
UK legal origin	1.313 (0.864)	5.812* (3.162)	-0.253 (1.522)	0.236 (6.453)	0.265** (0.109)	2.510 (5.735)	-6.764 (6.541)
Constant	94.431*** (1.380)	79.034*** (5.246)	0.580 (1.719)	81.059*** (9.356)	1.522*** (0.201)	69.115*** (11.581)	8.958 (7.146)
Observations	173	173	175	140	137	173	174
R^2	0.727	0.560	0.577	0.619	0.587	0.648	0.369

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table replicates the specifications of Table 2 of the main paper (in addition to column 6 of Table 1), but without controlling for Log GDP per capita. The column headings give the dependent variables for each of the columns. The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation and are based on the authors' calculations. The variable definitions and data sources for each of the variables are provided in Appendix B of the main paper.

Table 2: Child Survival: Decentralization, Linguistic Aggregation, Income & Ethnic Inequality, Segregation – No GDP

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Decentralization	Level 15	Level 2	10 × 10	Gini	Linguistic Inequality	Linguistic Segregation
Global ELF	-6.563*** (1.722)	-5.625*** (1.233)	-7.263*** (2.034)	-7.140*** (1.338)	-6.781*** (1.601)	-5.970*** (1.486)	-5.610*** (1.613)
Local-Global Compl.	20.263*** (7.020)	16.293*** (4.396)	17.752*** (6.237)	20.917*** (4.923)	18.618** (7.701)	16.925*** (5.151)	16.314*** (5.819)
Decentralization	1.289** (0.537)						
Decentralization × Global ELF	-1.252 (3.037)						
Decentralization × Local-Global	11.644 (14.019)						
Gini					0.035 (0.055)		
Linguistic Inequality						-2.349** (1.059)	
Linguistic Segregation							-2.330 (1.737)
Constant	93.868*** (1.651)	95.570*** (1.607)	93.808*** (1.421)	94.422*** (1.364)	93.580*** (2.562)	97.787*** (1.564)	95.327*** (1.481)
Baseline Controls (no GDP)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	151	173	173	173	124	165	162
R^2	0.762	0.722	0.694	0.729	0.736	0.740	0.736

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table replicates the specifications of Table 3 of the main paper, but without controlling for Log GDP per capita. The dependent variable is the child survival rate per 100 live births. The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation, except in columns 2 and 3 where they are measured at, respectively, levels 15 and 2 of aggregation. In column 1 decentralization is equal to 1 if country is a federal state, and 0 otherwise. In column 4, we use the Local-Global Ethnolinguistic Complementarity variable calculated at the 10 km × 10 km spatial resolution, instead of 5 km × 5 km. In column 5, we control for the income Gini coefficient, in column 6 we control for income inequality between linguistic groups, and in column 7 we control for linguistic segregation. All regressions control for the baseline controls except for Log GDP per capita: legal origins dummies, regional dummies and absolute latitude. The variable definitions and data sources for each of the variables are provided in Appendix B of the main paper.

Table 3: All Outcomes: Comprehensive Specification – No GDP

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Child Survival	Measles Immunization	Hospital Beds	Literacy Rate	Schooling	Improved Sanitation	Road Density
Global ELF	-5.679*** (1.292)	-26.868*** (4.778)	-1.659** (0.693)	-39.464*** (6.543)	-0.497*** (0.158)	-40.726*** (8.035)	-0.153 (2.310)
Local-Global Compl.	17.687*** (4.942)	64.003*** (20.553)	9.728*** (3.632)	136.152*** (24.430)	1.942*** (0.627)	122.724*** (37.099)	2.649 (13.934)
Log Population	-0.210 (0.173)	-0.930 (0.640)	-0.087 (0.124)	-0.757 (1.031)	-0.027 (0.019)	-1.079 (1.218)	-1.398*** (0.450)
Absolute Latitude	0.079*** (0.028)	0.122 (0.101)	0.095*** (0.023)	0.297 (0.194)	0.010*** (0.003)	0.416** (0.192)	0.261*** (0.097)
Soil Fertility	0.019 (0.743)	2.452 (2.682)	0.928 (0.614)	2.253 (5.041)	0.120 (0.085)	-7.725* (4.579)	-0.726 (2.236)
Roughness	3.111 (3.171)	-10.199 (11.815)	0.689 (2.136)	15.577 (16.080)	0.024 (0.319)	51.489** (21.740)	-18.606** (7.201)
Elevation	-1.232 (0.902)	0.639 (3.204)	-0.887 (0.591)	-0.094 (4.282)	-0.005 (0.089)	-15.879*** (5.774)	3.379** (1.698)
Island	-0.249 (1.006)	-7.559* (3.975)	0.738 (0.843)	-0.320 (7.956)	0.059 (0.106)	-8.183 (7.682)	4.377* (2.477)
Land Locked	-2.060*** (0.710)	-1.727 (2.480)	0.944** (0.474)	-2.195 (3.761)	-0.116 (0.080)	-1.528 (3.364)	-3.659** (1.432)
Constant	97.043*** (3.554)	104.261*** (12.905)	3.825 (3.592)	92.026*** (21.918)	2.237*** (0.428)	82.346*** (25.132)	30.587*** (10.709)
Baseline Controls (no GDP)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	149	149	149	123	127	147	149
R^2	0.811	0.627	0.654	0.647	0.625	0.741	0.454

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table replicates the specifications of Table 4 of the main paper, but without controlling for Log GDP per capita. The column headings give the dependent variables for each of the columns. The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation and are based on the authors' calculations. All regressions include the other baseline controls: legal origins dummies and regional dummies. The variable definitions and data sources for each of the variables are provided in Appendix B of the main paper.

Table 4: Other Public Goods

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Public Goods Index	Infant Survival	Electricity Access	Primary Enrollment	Skilled Births	Health Expenditure	Education Expenditure
Global ELF	-3.468*** (1.074)	-2.278*** (0.618)	-10.304 (7.746)	-17.423*** (4.488)	-29.295*** (6.203)	-0.530 (0.791)	-1.621* (0.949)
Local-Global Compl.	12.683*** (4.421)	6.444*** (2.333)	60.911** (28.378)	48.379*** (16.134)	116.071*** (21.666)	-3.657 (2.835)	1.133 (2.358)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	92	171	174	163	168	170	165
R^2	0.810	0.823	0.809	0.621	0.689	0.209	0.170
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	DPT Immunization	Hep B3 Immunization	Primary Completion	L. Sec. Completion	No Open Defecation	Basic Drinking Water	Basic Sanitation
Global ELF	-23.530*** (5.068)	-27.943*** (7.574)	-19.001*** (7.060)	-19.302*** (7.334)	-18.389** (7.210)	-8.724 (5.348)	-24.518*** (6.165)
Local-Global Compl.	65.089*** (16.625)	101.357*** (27.733)	56.670** (25.612)	98.256*** (26.089)	64.744*** (21.772)	29.501* (17.304)	78.071*** (23.211)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	171	158	160	161	174	174	174
R^2	0.573	0.271	0.630	0.741	0.476	0.747	0.813

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The specification is the one of Table 2 in the main paper. The column headings give the dependent variables for each of the columns. The dependent variables are all from the World Development Indicators (World Bank). Public Goods Index is the first principal component of the 7 primary dependent variables used in the paper along with the 13 additional dependent variables used in this table. The other dependent variables are defined as follows: Infant survival - survival of infants (less than one year of age) per 100 live births; Electricity Access - % of population with access to electricity; Primary Enrollment - adjusted net enrollment rate, primary (% of primary school age children); Skilled Births - births attended by skilled health staff (% of total); Health Expenditure - current health expenditure (% of GDP); Education Expenditure - total government expenditure on education (% of GDP); DPT Immunization - % of children ages 12-23 months immunized for DPT; Hep B3 Immunization - % of one-year-old children immunized for Hep B3; Primary Completion - total primary completion rate (% of relevant age group); L. Sec. Completion - total lower secondary completion rate (% of relevant age group); No Open Defecation - % of population not practicing open defecation; Basic Drinking Water - people using at least basic drinking water services (% of population); Basic Sanitation - people using at least basic sanitation services (% of population). The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation and are based on the authors' calculations. All columns include controls for regional and legal origin dummies, absolute latitude and Log GDP per capita. The control variable definitions and data sources are provided in Appendix B of the main paper.

Table 5: Measles Immunization: Decentralization, Linguistic Aggregation, Income & Ethnic Inequality, and Segregation

Dependent Variable: Measles Immunization	Decentralization	Level 15	Level 2	10 × 10	Gini	Linguistic Inequality	Linguistic Segregation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Global ELF	-20.439*** (6.471)	-18.926*** (3.924)	-28.838*** (7.032)	-24.831*** (4.494)	-29.571*** (4.283)	-24.146*** (4.740)	-22.215*** (6.248)
Local-Global Compl.	54.371** (23.481)	55.709*** (13.779)	70.830*** (19.035)	65.291*** (15.376)	76.864*** (17.635)	64.754*** (16.583)	61.225*** (21.767)
Decentralization	-2.500 (3.830)						
Decentralization × Global ELF	-12.720 (9.157)						
Decentralization × Local-Global	54.592 (52.730)						
Gini					0.185 (0.160)		
Linguistic Inequality						-3.390 (3.544)	
Linguistic Segregation							-2.987 (7.229)
Constant	67.381*** (10.467)	71.511*** (8.469)	60.564*** (9.015)	66.591*** (8.739)	71.178*** (9.086)	75.133*** (9.657)	67.178*** (9.211)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	149	171	171	171	124	165	161
R^2	0.576	0.555	0.540	0.571	0.641	0.573	0.570

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table replicates the specification of Table 3 in the main paper with Measles Immunization as its dependent variable. The column headings highlight the robustness test undertaken in each column. The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation, except in columns 2 and 3 where they are measured at, respectively, levels 15 and 2 of aggregation. Column 1 controls for decentralization and its interaction with Global ELF and the Local-Global Ethnolinguistic Complementarity variables. Decentralization is equal to 1 if country is a federal state, and 0 otherwise. In column 4, we use the Local-Global Ethnolinguistic Complementarity variable calculated at the 10 km × 10 km spatial resolution, instead of 5 km × 5 km. In column 5, we control for the income Gini coefficient, in column 6 we control for income inequality between linguistic groups, and in column 7 we control for linguistic segregation. All columns also control for Log GDP per capita, legal origins dummies, regional dummies and absolute latitude. The variable definitions and data sources for each of the variables are provided in Appendix B of the main paper.

Table 6: Hospital Beds: Decentralization, Linguistic Aggregation, Income & Ethnic Inequality, and Segregation

Dependent Variable: Hospital Beds	Decentralization	Level 15	Level 2	10 × 10	Gini	Linguistic Inequality	Linguistic Segregation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Global ELF	-0.936 (0.831)	-1.132** (0.535)	-0.440 (0.956)	-1.265** (0.612)	-1.623* (0.838)	-0.848 (0.679)	-1.452* (0.824)
Local-Global Compl.	9.376** (3.698)	3.804 (3.339)	5.117 (3.455)	7.159** (3.030)	10.403** (4.181)	6.249** (3.056)	7.729** (3.447)
Decentralization	0.422 (0.851)						
Decentralization × Global ELF	0.054 (1.481)						
Decentralization × Local-Global	-13.780 (8.325)						
Gini					-0.018 (0.030)		
Linguistic Inequality						-0.571 (0.641)	
Linguistic Segregation							0.185 (1.005)
Constant	-2.109 (2.199)	-1.625 (2.007)	-2.674 (1.969)	-2.023 (1.961)	-1.111 (2.250)	-1.336 (2.235)	-3.462* (1.866)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	149	173	173	173	124	165	162
R^2	0.630	0.593	0.592	0.598	0.665	0.607	0.590

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table replicates the specification of Table 3 in the main paper with Hospital Beds as its dependent variable. The column headings highlight the robustness test undertaken in each column. The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation, except in columns 2 and 3 where they are measured at, respectively, levels 15 and 2 of aggregation. Column 1 controls for decentralization and its interaction with Global ELF and the Local-Global Ethnolinguistic Complementarity variables. Decentralization is equal to 1 if country is a federal state, and 0 otherwise. In column 4, we use the Local-Global Ethnolinguistic Complementarity variable calculated at the 10 km × 10 km spatial resolution, instead of 5 km × 5 km. In column 5, we control for the income Gini coefficient, in column 6 we control for income inequality between linguistic groups, and in column 7 we control for linguistic segregation. All columns also control for Log GDP per capita, legal origins dummies, regional dummies and absolute latitude. The variable definitions and data sources for each of the variables are provided in Appendix B of the main paper.

Table 7: Literacy: Decentralization, Linguistic Aggregation, Income & Ethnic Inequality, and Segregation

Dependent Variable: Literacy	Decentralization	Level 15	Level 2	10 × 10	Gini	Linguistic Inequality	Linguistic Segregation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Global ELF	-30.380*** (7.644)	-18.012*** (4.566)	-32.886*** (9.561)	-31.770*** (5.903)	-35.335*** (7.003)	-35.779*** (6.171)	-28.705*** (7.825)
Local-Global Compl.	123.811*** (30.757)	54.846*** (19.654)	83.024*** (29.238)	102.715*** (21.674)	133.195*** (28.844)	110.042*** (22.752)	91.560*** (27.134)
Decentralization	3.880 (2.405)						
Decentralization × Global ELF	1.926 (11.600)						
Decentralization × Local-Global	-125.408** (60.047)						
Gini					0.233 (0.190)		
Linguistic Inequality						6.104 (3.950)	
Linguistic Segregation							-4.670 (9.476)
Constant	41.149*** (13.644)	42.087*** (11.031)	35.956*** (11.285)	45.831*** (10.836)	38.917*** (14.342)	44.389*** (11.179)	47.030*** (11.498)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	122	138	138	138	102	132	130
R ²	0.684	0.616	0.623	0.668	0.725	0.675	0.672

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table replicates the specification of Table 3 in the main paper with Literacy as its dependent variable. The column headings highlight the robustness test undertaken in each column. The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation, except in columns 2 and 3 where they are measured at, respectively, levels 15 and 2 of aggregation. Column 1 controls for decentralization and its interaction with Global ELF and the Local-Global Ethnolinguistic Complementarity variables. Decentralization is equal to 1 if country is a federal state, and 0 otherwise. In column 4, we use the Local-Global Ethnolinguistic Complementarity variable calculated at the 10 km × 10 km spatial resolution, instead of 5 km × 5 km. In column 5, we control for the income Gini coefficient, in column 6 we control for income inequality between linguistic groups, and in column 7 we control for linguistic segregation. All columns also control for Log GDP per capita, legal origins dummies, regional dummies and absolute latitude. The variable definitions and data sources for each of the variables are provided in Appendix B of the main paper.

Table 8: Schooling: Decentralization, Linguistic Aggregation, Income & Ethnic Inequality, and Segregation

Dependent Variable:	Decentralization	Level 15	Level 2	10 × 10	Gini	Linguistic Inequality	Linguistic Segregation
Schooling	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Global ELF	-0.288 (0.188)	-0.073 (0.101)	-0.280 (0.183)	-0.267* (0.136)	-0.365** (0.150)	-0.355*** (0.131)	-0.190 (0.205)
Local-Global Compl.	1.651*** (0.627)	0.717 (0.474)	1.122* (0.568)	1.517*** (0.498)	1.948*** (0.526)	1.556*** (0.490)	1.094 (0.690)
Decentralization	-0.007 (0.076)						
Decentralization × Global ELF	-0.019 (0.307)						
Decentralization × Local-Global	-1.121 (1.600)						
Gini					0.005 (0.004)		
Linguistic Inequality						0.076 (0.093)	
Linguistic Segregation							-0.159 (0.261)
Constant	0.589** (0.286)	0.601** (0.263)	0.575** (0.259)	0.664*** (0.252)	0.768*** (0.253)	0.756*** (0.247)	0.717*** (0.240)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	128	136	136	136	107	131	128
R^2	0.699	0.660	0.666	0.681	0.727	0.691	0.683

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table replicates the specification of Table 3 in the main paper with Schooling as its dependent variable. The column headings highlight the robustness test undertaken in each column. The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation, except in columns 2 and 3 where they are measured at, respectively, levels 15 and 2 of aggregation. Column 1 controls for decentralization and its interaction with Global ELF and the Local-Global Ethnolinguistic Complementarity variables. Decentralization is equal to 1 if country is a federal state, and 0 otherwise. In column 4, we use the Local-Global Ethnolinguistic Complementarity variable calculated at the 10 km × 10 km spatial resolution, instead of 5 km × 5 km. In column 5, we control for the income Gini coefficient, in column 6 we control for income inequality between linguistic groups, and in column 7 we control for linguistic segregation. All columns also control for Log GDP per capita, legal origins dummies, regional dummies and absolute latitude. The variable definitions and data sources for each of the variables are provided in Appendix B of the main paper.

Table 9: Improved Sanitation: Decentralization, Linguistic Aggregation, Income & Ethnic Inequality, and Segregation

Dependent Variable: Improved Sanitation	Decentralization	Level 15	Level 2	10 × 10	Gini	Linguistic Inequality	Linguistic Segregation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Global ELF	-29.748*** (7.716)	-25.439*** (5.872)	-26.511*** (8.593)	-24.844*** (6.034)	-33.665*** (7.171)	-22.458*** (6.950)	-20.268** (8.313)
Local-Global Compl.	112.937*** (33.567)	81.561*** (23.812)	73.915*** (28.237)	80.379*** (24.190)	121.690*** (34.573)	68.496** (27.442)	69.698** (31.378)
Decentralization	-0.609 (4.006)						
Decentralization × Global ELF	13.425 (11.886)						
Decentralization × Local-Global	-71.323 (82.897)						
Gini					-0.441 (0.304)		
Linguistic Inequality						-8.644* (4.926)	
Linguistic Segregation							-7.249 (9.897)
Constant	-1.719 (14.736)	-5.070 (13.373)	-24.022* (14.354)	-17.277 (14.070)	7.814 (14.950)	-0.312 (13.879)	-15.947 (14.508)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	147	171	171	171	123	163	161
R^2	0.821	0.795	0.772	0.783	0.809	0.797	0.779

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table replicates the specification of Table 3 in the main paper with Improved Sanitation as its dependent variable. The column headings highlight the robustness test undertaken in each column. The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation, except in columns 2 and 3 where they are measured at, respectively, levels 15 and 2 of aggregation. Column 1 controls for decentralization and its interaction with Global ELF and the Local-Global Ethnolinguistic Complementarity variables. Decentralization is equal to 1 if country is a federal state, and 0 otherwise. In column 4, we use the Local-Global Ethnolinguistic Complementarity variable calculated at the 10 km × 10 km spatial resolution, instead of 5 km × 5 km. In column 5, we control for the income Gini coefficient, in column 6 we control for income inequality between linguistic groups, and in column 7 we control for linguistic segregation. All columns also control for Log GDP per capita, legal origins dummies, regional dummies and absolute latitude. The variable definitions and data sources for each of the variables are provided in Appendix B of the main paper.

Table 10: Road Density: Decentralization, Linguistic Aggregation, Income & Ethnic Inequality, and Segregation

Dependent Variable: Road Density	Decentralization	Level 15	Level 2	10 × 10	Gini	Linguistic Inequality	Linguistic Segregation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Global ELF	1.229 (1.907)	1.753 (1.745)	1.113 (2.026)	0.768 (1.690)	0.697 (2.459)	-1.627 (2.229)	-1.571 (2.310)
Local-Global Compl.	8.521 (15.100)	-2.685 (11.654)	6.614 (12.029)	7.924 (11.971)	8.814 (17.248)	13.005 (12.146)	12.013 (12.865)
Decentralization	5.542 (4.165)						
Decentralization × Global ELF	-4.704 (6.369)						
Decentralization × Local-Global	-73.391 (52.563)						
Gini					0.037 (0.116)		
Linguistic Inequality						5.158* (2.709)	
Linguistic Segregation							3.635 (2.619)
Constant	0.831 (9.928)	-1.800 (9.673)	-0.351 (9.327)	-0.387 (9.522)	-5.091 (11.144)	-6.404 (10.188)	-6.191 (10.660)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	149	172	172	172	124	165	162
R^2	0.470	0.457	0.460	0.460	0.483	0.479	0.428

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table replicates the specification of Table 3 in the main paper with Road Density as its dependent variable. The column headings highlight the robustness test undertaken in each column. The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation, except in columns 2 and 3 where they are measured at, respectively, levels 15 and 2 of aggregation. Column 1 controls for decentralization and its interaction with Global ELF and the Local-Global Ethnolinguistic Complementarity variables. Decentralization is equal to 1 if country is a federal state, and 0 otherwise. In column 4, we use the Local-Global Ethnolinguistic Complementarity variable calculated at the 10 km × 10 km spatial resolution, instead of 5 km × 5 km. In column 5, we control for the income Gini coefficient, in column 6 we control for income inequality between linguistic groups, and in column 7 we control for linguistic segregation. All columns also control for Log GDP per capita, legal origins dummies, regional dummies and absolute latitude. The variable definitions and data sources for each of the variables are provided in Appendix B of the main paper.

Table 11: Multicollinearity: Breaking Local ELF into two parts - Child Survival

Dependent Variable	(1)	(2)	(3)	(4)	(5)
Child Survival					
Global ELF	-4.257*** (1.072)	-3.483** (1.456)	-4.257*** (1.072)	-3.483** (1.456)	-3.483** (1.456)
Local-Global Compl.	12.486*** (3.891)	18.351** (8.596)			
Local ELF		-3.157 (4.214)			-3.157 (4.214)
Local-Global Component			5.116*** (1.594)	4.361** (1.960)	7.518** (3.522)
Residual Local ELF				-3.157 (4.214)	
Constant	82.156*** (1.746)	82.233*** (1.755)	82.052*** (1.750)	82.080*** (1.749)	82.080*** (1.749)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Observations	171	171	171	171	171
R^2	0.825	0.826	0.825	0.826	0.826

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is child survival rate per 100 live births. The Global ELF, Local ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation and are based on the authors' calculations. Column 1 gives the baseline specification (column 6 of Table 1 in main paper) that controls for legal origins, Log GDP per capita and region-specific dummies. In column 2 we introduce a control for Local ELF. For the specifications in columns 3-5, we break the Local ELF variable into two orthogonal components by regressing Local ELF on the Local-Global Ethnolinguistic Complementarity variable. The first component - Local-Global Component - gives the part of the Local ELF variable that is solely correlated to (predicted by) our Local-Global Ethnolinguistic Complementarity variable. The second component - Residual Local ELF - gives us the residual variation in Local ELF which is orthogonal to our Local-Global Ethnolinguistic Complementarity. In column 3 we only control for the Local-Global Component of Local ELF. In column 4 we add a control for the residual component of Local ELF. In column 5 we control for Local ELF and the Local-Global Component. The variable definitions and data sources for each of the other variables are provided in Appendix B of the main paper. In terms of interpretation, column 2 shows that when jointly controlling for Local-Global Complementarity and Local ELF, the former is statistically significant, whereas the latter is not. Columns 3-5 show that when splitting up Local ELF into a component that is predicted by Local-Global Complementarity and an orthogonal component, only the one correlated to Local-Global Complementarity is statistically significant.

Table 12: Multicollinearity: Breaking Local ELF into two parts - All Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Child Survival	Measles Immunization	Hospital Beds	Literacy Rate	Schooling	Improved Sanitation	Road Density
Global ELF	-3.483** (1.456)	-27.348*** (6.092)	-2.694*** (0.867)	-30.643*** (10.187)	-0.067 (0.194)	-34.076*** (8.038)	4.020 (3.130)
Local-Global Component	4.361** (1.960)	29.580*** (7.659)	4.301*** (1.356)	40.139*** (12.833)	0.389 (0.254)	42.342*** (11.080)	-0.599 (5.290)
Residual Local ELF	-3.157 (4.214)	9.433 (15.318)	5.897** (2.657)	-4.116 (29.997)	-0.776 (0.655)	36.712* (21.337)	-11.715 (10.599)
Constant	82.080*** (1.749)	65.957*** (8.856)	-2.140 (1.961)	44.993*** (10.987)	0.641** (0.253)	-18.304 (14.131)	-0.598 (9.510)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	171	171	173	138	136	171	172
R^2	0.826	0.573	0.604	0.665	0.680	0.786	0.462

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The specification is the one of column 4 in Table 11. The column headings give the dependent variables for each of the columns. Local-Global Component and Residual Local ELF are the predicted part and the residual of a regression of Local ELF on Local-Global Ethnolinguistic Complementarity. The variable definitions and data sources for each of the variables are provided in Appendix B of the main paper. In terms of interpretation, the table shows that when splitting up Local ELF into a component that is predicted by Local-Global Complementarity and an orthogonal component, the one correlated to Local-Global Complementarity is in general the statistically more significant one.

Table 13: Alternative Local-Global Ethnolinguistic Complementarity Index

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Child Survival	Measles Immunization	Hospital Beds	Literacy Rate	Schooling	Improved Sanitation	Road Density
Global ELF	-4.301*** (1.483)	-30.503*** (6.461)	-2.802*** (0.862)	-35.658*** (9.780)	-0.198 (0.192)	-36.814*** (8.182)	3.188 (2.944)
Alternative Local-Global Compl.	5.100 (3.190)	41.285*** (12.900)	6.986*** (2.158)	49.757** (20.913)	0.462 (0.443)	64.255*** (17.713)	-3.321 (7.721)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	171	171	173	138	136	171	172
R^2	0.817	0.556	0.604	0.635	0.658	0.783	0.458

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The column headings give the dependent variables for each of the columns. The Global ELF and the Alternative Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation and are based on the authors' calculations. The Alternative Local-Global Ethnolinguistic Complementarity is defined as $\sum_{\ell} s_{\ell} \sum_i s_{\ell i} (1 - s_{\ell i})(1 - s_i)$. The specification is the same as Table 2 in the main paper, so all regressions control for legal origins, region-specific dummy variables and log GDP per capita. The control variable definitions and data sources are provided in Appendix B of the main paper.

Table 14: Further Region-Specific Robustness

	Drop sub- Sahara (1)	Drop East S.E. Asia (2)	Drop Latin America (3)	Colonial Origin (4)	WB Regions (5)
Dependent Variable: Child Survival					
Global ELF	-2.819*** (0.925)	-4.635*** (1.189)	-4.874*** (1.169)	-4.030*** (1.081)	-4.051*** (1.154)
Local-Global Compl.	7.006** (2.766)	12.672*** (4.067)	15.959*** (4.520)	10.718** (4.435)	11.239*** (4.210)
Observations	125	159	139	175	171
R^2	0.687	0.832	0.826	0.826	0.826
Dependent Variable: Measles Immunization					
Global ELF	-18.037*** (5.286)	-24.683*** (4.960)	-28.464*** (4.793)	-21.738*** (4.590)	-23.930*** (4.711)
Local-Global Compl.	54.473*** (14.828)	63.674*** (16.415)	78.537*** (18.180)	49.424*** (17.106)	59.992*** (16.424)
Observations	125	159	139	175	171
R^2	0.301	0.575	0.586	0.611	0.571
Dependent Variable: Hospital Beds					
Global ELF	-1.718* (1.022)	-1.752*** (0.582)	-1.205* (0.688)	-0.472 (0.436)	-1.240** (0.519)
Local-Global Compl.	8.171** (3.546)	8.441*** (2.893)	7.582** (3.577)	-3.037 (2.076)	6.037** (2.604)
Observations	127	161	139	177	173
R^2	0.543	0.613	0.623	0.739	0.666
Dependent Variable: Literacy					
Global ELF	-19.895** (7.819)	-34.242*** (6.519)	-34.821*** (6.551)	-27.256*** (5.111)	-32.257*** (5.391)
Local-Global Compl.	87.961*** (21.838)	107.781*** (23.560)	103.771*** (26.442)	59.354*** (19.395)	95.313*** (19.076)
Observations	94	128	112	141	138
R^2	0.470	0.675	0.672	0.753	0.744
Dependent Variable: Schooling					
Global ELF	-0.231 (0.139)	-0.294* (0.155)	-0.290* (0.151)	-0.123 (0.126)	-0.263* (0.137)
Local-Global Compl.	1.447*** (0.469)	1.424*** (0.532)	1.508*** (0.532)	0.090 (0.528)	1.296*** (0.455)
Observations	104	124	112	138	136
R^2	0.583	0.681	0.688	0.708	0.729
Dependent Variable: Improved Sanitation					
Global ELF	-24.646*** (8.631)	-25.338*** (6.475)	-28.471*** (6.330)	-20.466*** (6.057)	-24.384*** (5.970)
Local-Global Compl.	72.968*** (26.602)	78.682*** (24.892)	96.894*** (30.048)	48.711* (24.822)	72.159*** (24.039)
Observations	125	159	137	174	171
R^2	0.617	0.791	0.799	0.795	0.810
Dependent Variable: Road Density					
Global ELF	1.075 (2.457)	0.463 (1.971)	1.011 (1.882)	1.540 (1.860)	0.075 (1.318)
Local-Global Compl.	6.285 (14.160)	6.409 (12.564)	6.639 (15.285)	1.290 (12.977)	11.002 (11.616)
Observations	126	160	139	175	172
R^2	0.448	0.474	0.473	0.385	0.501

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation and are based on the authors' calculations. All columns control for baseline regional dummies and Log GDP per capita. Columns 1, 2, 3 and 5 control for legal origins. Column 1 drops countries from sub-Saharan Africa, column 2 drops countries from East and South East Asia, column 3 drops countries from Latin America and the Caribbean, column 4 replaces legal origin by colonial origin, column 5 replaces the three regional dummies of the baseline by six regional dummies commonly used by the World Bank. The colonial origins variable gives us the country from which a country became independent (Source: CIA World Factbook, 2001). Other variable definitions and data sources are provided in Appendix B of the main paper.

Table 15: Heterogeneity by Traditional Occupational Specialization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Child Survival	Measles Immunization	Hospital Beds	Literacy Rate	Schooling	Improved Sanitation	Road Density
Global ELF	-3.803*** (1.235)	-21.299*** (4.886)	-0.861 (0.652)	-32.528*** (6.463)	-0.239 (0.149)	-19.535*** (6.919)	1.055 (2.263)
Local-Global Compl.	11.811*** (3.814)	61.441*** (14.318)	6.367** (2.890)	98.525*** (20.403)	1.319*** (0.465)	73.590*** (22.655)	5.657 (12.983)
Specialization	-0.362 (0.353)	-2.792 (1.727)	-0.391** (0.195)	-0.534 (1.451)	-0.020 (0.045)	-4.630** (2.056)	0.158 (0.918)
Local-Global \times Specialization	3.130 (2.382)	21.564* (11.318)	4.575** (2.196)	24.097** (10.340)	0.385 (0.302)	43.187*** (14.164)	-2.605 (7.404)
Observations	171	171	173	138	136	171	172
R^2	0.827	0.583	0.608	0.676	0.680	0.793	0.459

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The column headings give the dependent variables for each of the columns. The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation and are based on the authors' calculations. The Specialization variable gives us an index of ethnic specialization based on historical occupational specialization data from Murdock's Ethnographic Atlas (extended by Giuliano and Nunn, 2018). These data were originally available at ethnicity level which were matched to languages from the Ethnologue. Specialization is measured as historical occupational specialization based on historical subsistence dependence on each of the following 5 occupations: Gathering, Hunting, Fishing, Animal Husbandry and Agriculture. For a particular country, to measure the specialization of different ethnic groups i across sectors q , we define the following variables: x_{iq} : share of ethnic group i employed in sector q ; x_q : share of country employed in sector q ; pop_i : share of ethnic group i in the total population of the country. Then the specialization for ethnic group i (proposed by Krugman, 1991) is defined as $spec_i = \sum_q |x_{iq} - x_q|$, and the specialization of the country is defined as: Specialization = $pop_i spec_i$. Finally, we standardize the Specialization Index to have mean 0 and standard deviation 1 for ease of interpretation of the coefficients. The other control variable definitions and data sources are provided in Appendix B of the main paper.

Table 16: Child Survival by Decades

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	1960s	1970s	1980s	1990s	200s0	2010s	1990-2010
Panel A: WDI GDP Data							
Global ELF	-12.135*** (2.653)	-9.119*** (2.353)	-7.529*** (1.614)	-4.640*** (1.274)	-3.661*** (0.939)	-3.080*** (0.793)	-4.257*** (1.072)
Local-Global Compl.	22.348* (12.528)	15.881 (9.955)	14.107* (7.597)	13.736*** (4.361)	10.879*** (3.523)	8.665*** (2.811)	12.486*** (3.891)
Observations	97	113	145	170	172	171	171
R^2	0.773	0.747	0.779	0.811	0.828	0.795	0.825
Panel B: PWT GDP Data							
Global ELF	-13.971*** (2.930)	-8.684*** (2.192)	-6.936*** (1.639)	-4.703*** (1.312)	-3.412*** (1.011)	-2.939*** (0.877)	-4.257*** (1.072)
Local-Global Compl.	40.996*** (12.849)	24.947** (10.239)	17.030** (7.204)	12.712*** (4.253)	9.746*** (3.598)	7.835*** (3.001)	12.486*** (3.891)
Observations	101	135	142	163	163	163	171
R^2	0.758	0.720	0.792	0.831	0.843	0.802	0.825
Panel C: Not controlling for GDP							
Global ELF	-13.111*** (3.366)	-12.208*** (2.701)	-10.806*** (2.039)	-8.276*** (1.675)	-5.929*** (1.157)	-4.404*** (0.852)	-7.089*** (1.362)
Local-Global Compl.	53.592*** (10.600)	42.534*** (8.406)	32.244*** (6.849)	24.198*** (5.888)	17.257*** (4.370)	12.497*** (3.254)	20.512*** (4.996)
Observations	127	155	173	173	173	173	173
R^2	0.575	0.587	0.650	0.689	0.749	0.727	0.727

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is child survival rate per 100 live births. The column headings give the decade(s) for which child survival has been measured (based on the WDI data). 1960s refers to all years (with available data) between 1960 and 1969 (including 1960), 1970s refers to all years (with available data) between 1970 and 1979 (including 1970), and so on. 2010s refers to 2010 until the most recent year (2016 for most countries), and 1990-2010 refers to all years between 1990 and 2010 (including 1990 and 2010). The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation and are based on the authors' calculations. All specifications control for legal origin and region-specific dummy variables. Panel A controls for Log GDP per capita from the World Development Indicators database. Panel B instead controls for Log GDP per capita from the Penn World Tables database. Panel C does not include any controls for GDP. The control variable definitions and data sources are provided in Appendix B of the main paper.

Table 17: Heterogeneity by Motorization Rate

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Child Survival	Measles Immunization	Hospital Beds	Literacy Rate	Schooling	Improved Sanitation	Road Density
Global ELF	-3.185** (1.328)	-21.500*** (5.463)	-0.986 (0.691)	-20.027** (7.909)	-0.096 (0.170)	-24.224*** (7.295)	1.485 (2.414)
Local-Global Compl.	12.541** (5.122)	63.889*** (20.221)	9.165*** (3.396)	89.273*** (24.834)	1.116** (0.552)	96.184*** (28.367)	2.248 (11.953)
Motorization Rate	0.436 (0.392)	-1.523 (2.102)	1.328*** (0.369)	4.775** (2.307)	0.113*** (0.041)	4.442** (2.075)	3.983** (1.974)
Local-Global \times Motorization	-7.881** (3.723)	-17.168 (16.374)	-6.231* (3.377)	-55.939*** (19.656)	-0.753** (0.333)	-57.151** (22.425)	14.701 (20.395)
Observations	129	129	131	105	115	128	130
R^2	0.853	0.541	0.667	0.671	0.705	0.842	0.535

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The column headings give the dependent variables for each of the columns. The Global ELF and the Local-Global Ethnolinguistic Complementarity variables are measured at level 5 of aggregation and are based on the authors' calculations. The Motorization variable gives us the standardized motorization rate in the country. The motorization rate variable is defined as the number of vehicles per 1,000 inhabitants in the country in the year 2015. These data available for 141 countries are provided by the International Organization of Motor Vehicle Manufacturers (OICA) and were downloaded from: <http://www.oica.net/category/vehicles-in-use/>. All specifications control for legal origin, region-specific dummy variables and Log GDP per capita. The control variable definitions and data sources are provided in Appendix B of the main paper.